International Polar-Orbit Processing Package Framework for Earth Remote Sensing Science Data Processing

Patrick L. Coronado[1]*, John Overton[2] and Kelvin W. Brentzel[1]

[1] Direct Readout Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2] NPOESS Integrated Program Office, Silver Spring MD, USA

Introduction

The NPOESS Integrated Program Office (IPO) is chartered to develop a direct readout processing package to produce data products from earth observation satellites starting with NASA's EOS missions and evolving to the NPOESS instrument suites. This development was built upon a technology suite comprising the NPP In-Situ Ground system (NISGS). Applying these NISGS technologies has enabled the creation of a system framework used in the International Polar Orbiter Processing Package (IPOPP). IPOPP is the primary software package that will allow the direct readout community to process, visualize, and evaluate Earth science data from the MODIS instrument on Aqua and Terra spacecraft, and future National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP) and NPOESS instrument suite. IPOPP's primary role is to maximize the utility of direct broadcast data by generating regional real-time products for insertion into a decision-making infrastructure. IPOPP development represents collaboration between the DRL, the NOAA Integrated Program Office (IPO), and the University of Wisconsin-Madison (UW) – providing direct readout atmosphere Environmental Data Record (EDR) algorithms. The DRL and its collaborators also work closely with the direct readout community to ensure that IPOPP functionality meet user needs (see Figure 1). IPOPP will be:

- Freely available
- Portable to Linux x86 platforms
- Efficient to run on modest hardware
- Simple to install and easy to use
- Able to ingest and process Direct Broadcast overpasses of arbitrary size

Able to produce core and regional value-added Environmental Data Record (EDR) products

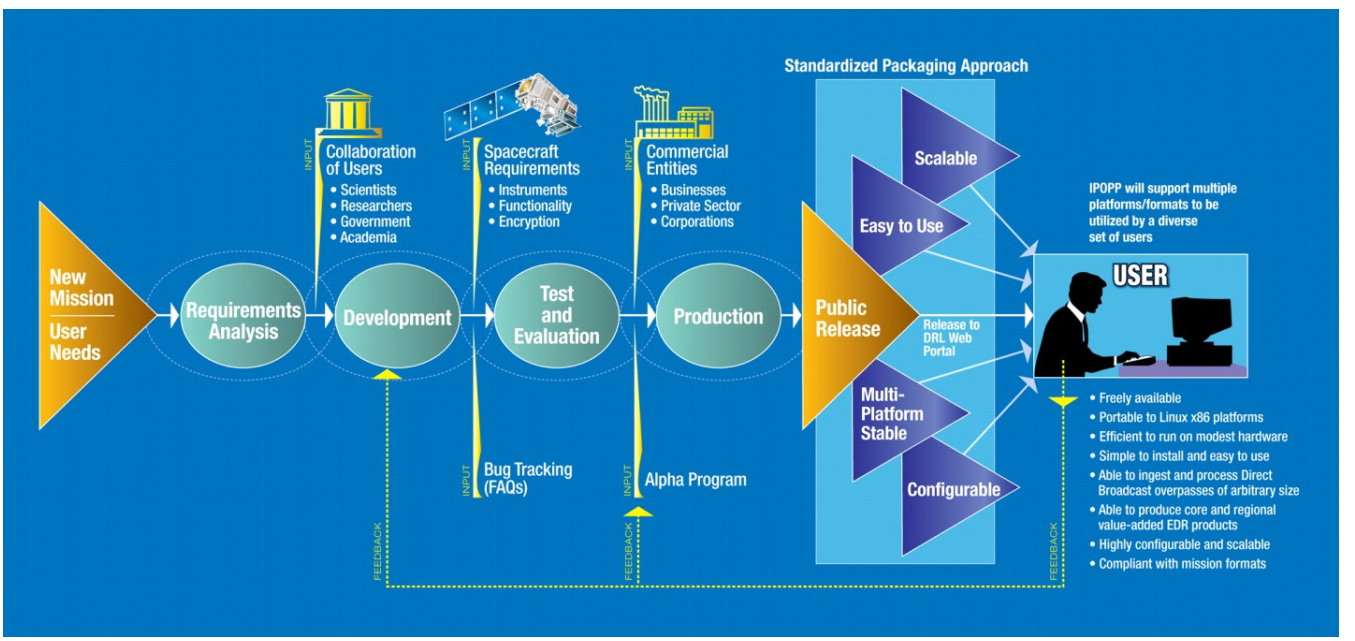

Figure 1 IPOPP Development Process

*Corresponding :patrick.coronado@nasa.gov

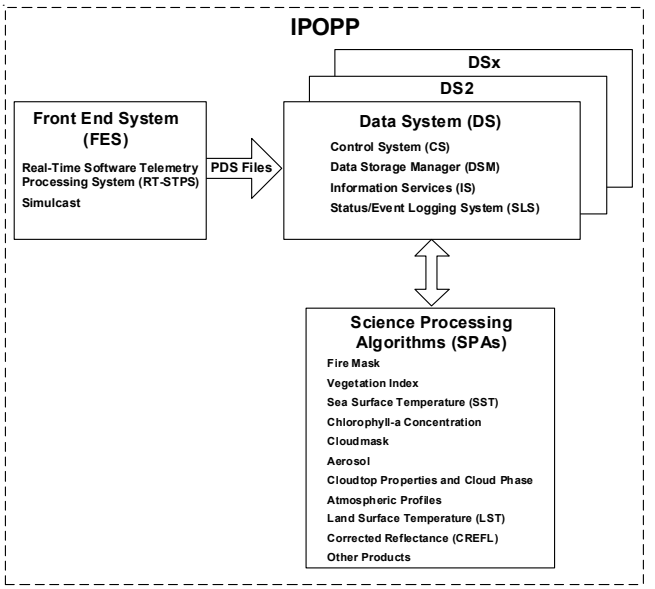

Figure 2 Major IPOPP Components

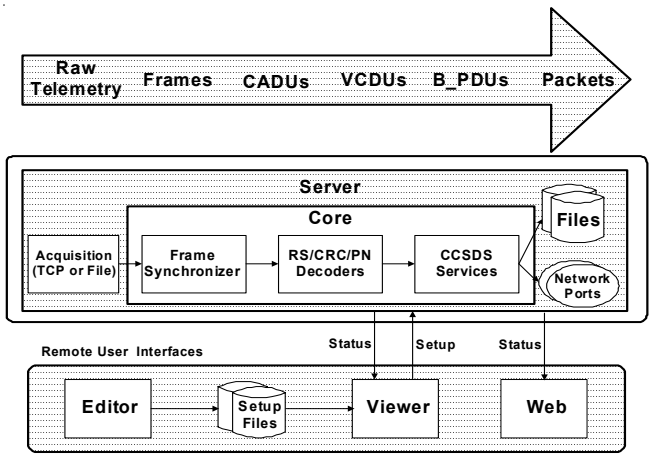

Figure 3 RT-STPS Architecture

IPOPP Overview

The IPOPP framework includes the Front End System (FES) and the Data System (DS). The FES contains the Real-time Software Telemetry Processing System (RT-STPS) and Simulcast. The DS contains the Control System (CS), the Data Storage Manager (DSM), the Information Services (IS), and the Status/Event Logging System (SLS). A third component of the IPOPP system are the EDRs themselves which are managed by a Science Processing Algorithm (SPA) wrapper. The SPA wrapper also provides extensible functionality and serves as a process catalyst for the DS. See Figure 2, "Major IPOPP Components." As depicted in Figure 2, IPOPP is extensible, meaning that the user can have any number of Data Systems to process any number of SPAs.

Front End System (FES)

The FES contains the Real-time Software Telemetry Processing System (RT-STPS) and Simulcast. The RT-STPS Server and RT-STPS Batch Processor are Java applications that ingest raw telemetry data and produce Level 0 products – similar to Raw Data Records (RDRs), including sorted Consultative Committee for Space Data Systems (CCSDS) packets and Virtual Channel Data Units (VCDUs). The RT-STPS Batch Processor ingests a single data file as specified in the configuration file. The RT-STPS Server runs continuously. The RT-STPS Server and Batch Processor can each send data across Transmission Control Protocol (TCP)/ Internet Protocol (IP) ports, to files, or to both simultaneously, depending on the configuration file.

The RT-STPS package includes three utilities: the Viewer, the Server, and the Sender. The RT-STPS Viewer displays the progress of the RT-STPS Server as it runs, and loads and unloads the Server configuration files. The RT-STPS Sender copies a raw data file to the RT-STPS Server.

RT-STPS Architecture is depicted in Figure 3.

Simulcast is a real-time Java application that allows users to select and view quicklook instrument data from multiple missions and spacecraft. Simulcast provides real-time geolocation and pseudo-calibration, and projects data on Mercator and Polar maps. Simulcast can replay recent satellite passes, export displayed images to JPEG format, and save replayed passes to AVI/Quicktime movies.

The Simulcast Client consists of the Simulcast Viewer and the Simulcast Console. The Viewer is used to display data from satellite passes (currently MODIS data from Aqua and Terra). The Console is used to display and control administrative information. The Simulcast Client is depicted in Figure 4.

Data System (DS)

The DS contains the Control System (CS), the Data Storage Manager (DSM), the Information Services (IS), and the Status/Event Logging System (SLS). The CS assembles the requisite input resources and schedules the execution of the Science Processing Algorithms (SPAs). SPAs generate the Level 1 - similar to thScience Data Records (SDRs) and Level 2 end products – same as EDRs. The DSM maintains

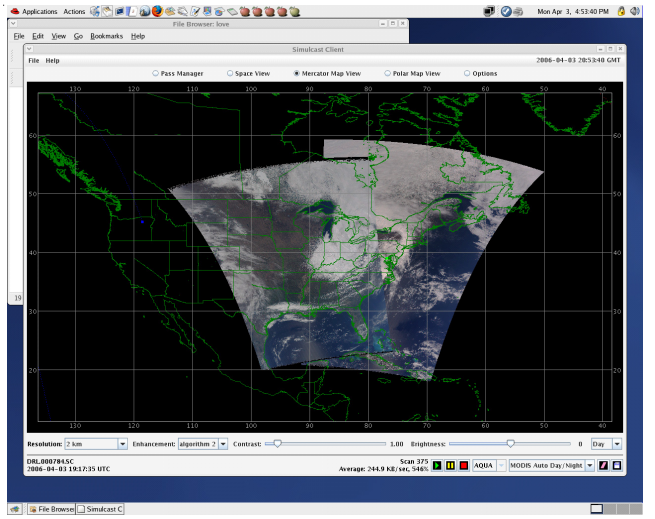

Figure 4 Simulcast Client (Mercator Map View)

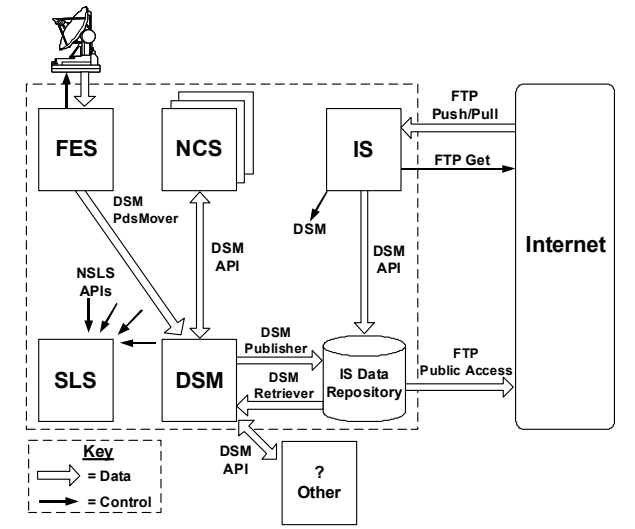

Figure 5 IPOPP Framework Data Flow

a MySQL database describing the location of data files and products. DSM agents move data among the several IPOPP components and store all products in the IS Data Repository. The IS maintains a publicly accessible static subdirectory, the IS Data Repository, where products and data files are stored. The IS retrieves ancillary data files from remote locations and places them into this subdirectory. The SLS manages and displays messages logged by the IPOPP components.

The FES places PDS and CSR files in a specified location where the DSM PdsMover agent discovers and copies them to the IS Repository while registering the files with the DSM. Using File Transfer Protocol (FTP) or Hypertext Transfer Protocol (HTTP) as defined in a configuration file, the IS retrieves ancillary data files from remote Internet sites, stores these data at specified locations in the IS Repository, and registers them with the DSM. The IS is initially configured to retrieve ancillary data from the DRL's IS Data Repository; but this can be user defined. Ancillary data may also be pushed to the repository from remote sites using FTP. Each CS instance controls an SPA and is called a station. A CS station requests the requisite input products from the DSM and waits for the DSM Retriever agent to make the input products available before executing the SPA. The resulting output products are registered with the DSM. The DSM Publisher agent copies registered products to the IS Data Repository, where they are publicly accessible and available for subsequent retrieval by the DSM Retriever agent for inputs to successor SPAs. IPOPP framework data flow is depicted in Figure 5.

Science Processing Algorithms (SPAs)

The DRL has developed an algorithm wrapper that can be applied to virtually any processing algorithm so that it may be integrated easily into the IPOPP, as well as run in a standalone environment. Investigators have typically written processing algorithms to create Earth science data products in a variety of programming languages, with little regard for common command and execution interfaces. As a result, integrating these algorithms into previous direct readout processing packages has been a painstaking and labor-intensive exercise, performed on a case-by-case basis. No standardized packaging technique existed that would allow algorithms to be integrated easily into processing packages prior to the IPOPP, much less in a standalone environment, across platforms.

The algorithm wrapper part of the framework is key to the success of the IPOPP. Once an algorithm has been packaged with this wrapper, it is referred to as a Science Processing Algorithm (SPA). The SPA wrapper process is depicted in Figure 6. Algorithm wrappers provide a common command and execution interface to encapsulate multi-discipline, multi-mission SPAs. The wrapper also provides a structured, standardized technique for packaging new or updated algorithms with minimal effort. The DRL's Modular Components Approach (MCA) ensures that SPAs will function in a standalone environment, across platforms, to serve the needs of the broad direct readout community. See Figure 7. All of the processing algorithms and processing packages released by the DRL conform to this common Application Programming Interface (API).

Alpha Test Program

The IPOPP development approach is driven by user needs. The DRL has launched an Alpha Test Program to deliver evolving IPOPP technologies for evaluation and feedback.

The DRL distributes DVDs—loaded with IPOPP technologies, SPAs, and a User's Guide—to Alpha testers, who then install, configure and operate the technologies on

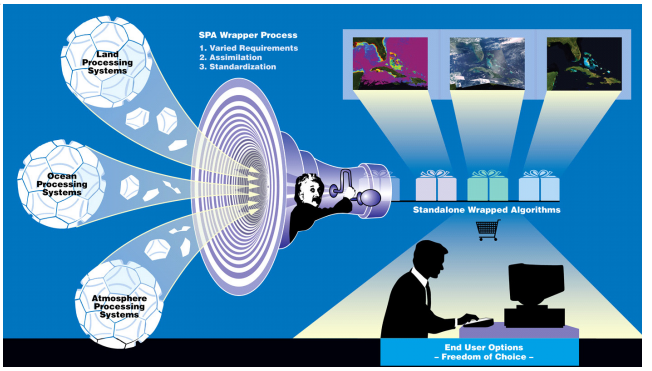

Figure 6 SPA Wrapper Process

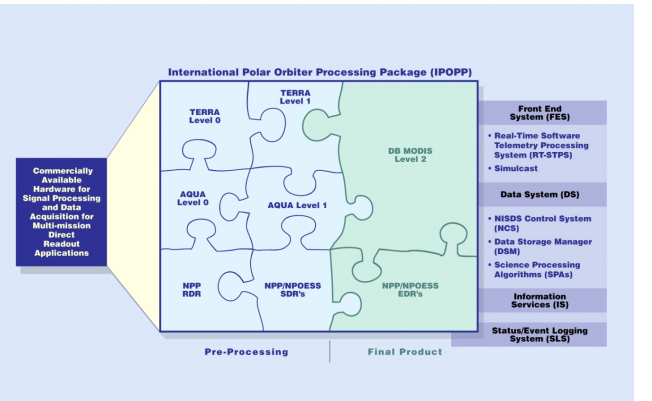

Figure 7 Modular Components Approach

their own hardware. DRL developers are available to provide technical support if necessary.

Alpha testers acquire the ability to create Land, Ocean, and Atmosphere products in real time. In return, Alpha testers provide the DRL with valuable feedback as to how the utility of IPOPP technologies might be enhanced prior to public release. The DRL will also gain insight into how well its SPA packaging techniques work across platforms. At the time of this writing, the DRL plans to add additional Alpha test sites with subsequent releases.

Summary

IPOPP constitutes a paradigm shift in how the United States Government provides it's internally-developed remote sensing science algorithms to the public. It provides for higher utility of NASA's Earth science data by directly providing technologies for decision-support infrastructure, and enabling solutions for real-time and regional applications. At the same time providing continuity among Earth remote sensing missions to minimize end-user impact and contribute to regional real-time application support systems.

At the core of this paradigm is the framework consisting of computer science technologies that provide for ubiquitous utility and transportability of Earth science data and its derivative products; with inherent modularity, scalability, portability & extensibility of direct readout tools for real-time data processing. And finally, with the underlying strength provided by the promotion of standards in pre-processing sub-systems, SPAs interfaces, visualization and real-time processing systems.

This paradigm, along with the IPOPP framework and its SPAs, underpins NASA's primary vision for DB, the unification of the global direct readout community for data democratization and understanding of Earth processes as a system.

References

Coronado P, Brentzel K, Harris C (2003a) NPP In-Situ Ground System: Bridging Technologies from EOS to NPOESS. International Telemetry Conference. November, Las Vegas, NV

Coronado P, Campbell W (1997) Applying Information technologies to Facilitate Information Access and Regional Development. Tech 2007 NASA Annual Conference

NASA Direct Readout and Its Polar Orbiting Satellites